



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 21, 2024 – 07:14 AM EDT

PDB ID : 1YLO
Title : Crystal Structure of Protein of Unknown Function (Possible Aminopeptidase)
S2589 from Shigella flexneri 2a str. 2457T
Authors : Nocek, B.P.; Midwest Center for Structural Genomics (MCSG)
Deposited on : 2005-01-19
Resolution : 2.15 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

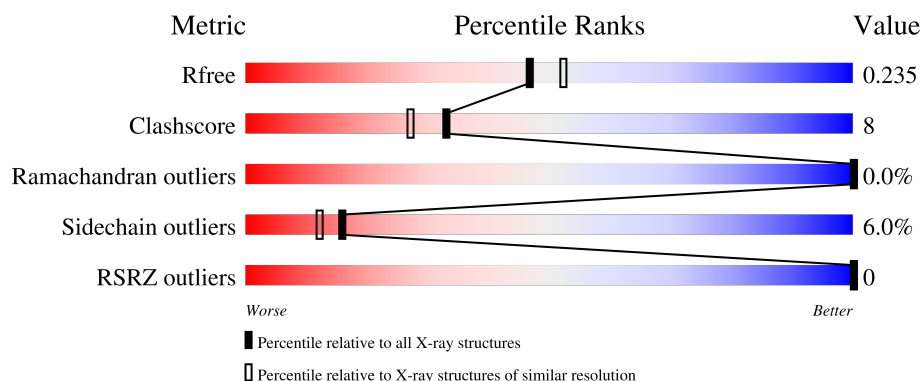
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.




Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1881 (2.16-2.16)
Clashscore	180529	2047 (2.16-2.16)
Ramachandran outliers	177936	2027 (2.16-2.16)
Sidechain outliers	177891	2026 (2.16-2.16)
RSRZ outliers	164620	1882 (2.16-2.16)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	348	
1	B	348	
1	C	348	
1	D	348	
1	E	348	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	348	 A horizontal bar chart showing the quality of the chain. The bar is divided into three segments: a long green segment representing 84%, a shorter yellow segment representing 13%, and a very short red segment representing 3%. The text '84%' is centered under the green segment, and '13%' is centered under the yellow segment. To the right of the red segment are two small black dots '••'.

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 16577 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called hypothetical protein SF2450.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	346	Total	C	N	O	S	Se	0	4	0
			2648	1647	475	508	6	12			
1	B	346	Total	C	N	O	S	Se	0	4	0
			2648	1647	475	508	6	12			
1	C	346	Total	C	N	O	S	Se	0	4	0
			2648	1647	475	508	6	12			
1	D	346	Total	C	N	O	S	Se	0	4	0
			2648	1647	475	508	6	12			
1	E	346	Total	C	N	O	S	Se	0	4	0
			2648	1647	475	508	6	12			
1	F	346	Total	C	N	O	S	Se	0	4	0
			2648	1647	475	508	6	12			

There are 90 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	cloning artifact	UNP Q83K87
A	-1	ASN	-	cloning artifact	UNP Q83K87
A	0	ALA	-	cloning artifact	UNP Q83K87
A	1	MSE	MET	modified residue	UNP Q83K87
A	58	MSE	MET	modified residue	UNP Q83K87
A	63	MSE	MET	modified residue	UNP Q83K87
A	69	MSE	MET	modified residue	UNP Q83K87
A	89	MSE	MET	modified residue	UNP Q83K87
A	123	MSE	MET	modified residue	UNP Q83K87
A	136	MSE	MET	modified residue	UNP Q83K87
A	161	MSE	MET	modified residue	UNP Q83K87
A	244	MSE	MET	modified residue	UNP Q83K87
A	277	MSE	MET	modified residue	UNP Q83K87
A	300	MSE	MET	modified residue	UNP Q83K87
A	322	MSE	MET	modified residue	UNP Q83K87
B	-2	SER	-	cloning artifact	UNP Q83K87
B	-1	ASN	-	cloning artifact	UNP Q83K87

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	ALA	-	cloning artifact	UNP Q83K87
B	1	MSE	MET	modified residue	UNP Q83K87
B	58	MSE	MET	modified residue	UNP Q83K87
B	63	MSE	MET	modified residue	UNP Q83K87
B	69	MSE	MET	modified residue	UNP Q83K87
B	89	MSE	MET	modified residue	UNP Q83K87
B	123	MSE	MET	modified residue	UNP Q83K87
B	136	MSE	MET	modified residue	UNP Q83K87
B	161	MSE	MET	modified residue	UNP Q83K87
B	244	MSE	MET	modified residue	UNP Q83K87
B	277	MSE	MET	modified residue	UNP Q83K87
B	300	MSE	MET	modified residue	UNP Q83K87
B	322	MSE	MET	modified residue	UNP Q83K87
C	-2	SER	-	cloning artifact	UNP Q83K87
C	-1	ASN	-	cloning artifact	UNP Q83K87
C	0	ALA	-	cloning artifact	UNP Q83K87
C	1	MSE	MET	modified residue	UNP Q83K87
C	58	MSE	MET	modified residue	UNP Q83K87
C	63	MSE	MET	modified residue	UNP Q83K87
C	69	MSE	MET	modified residue	UNP Q83K87
C	89	MSE	MET	modified residue	UNP Q83K87
C	123	MSE	MET	modified residue	UNP Q83K87
C	136	MSE	MET	modified residue	UNP Q83K87
C	161	MSE	MET	modified residue	UNP Q83K87
C	244	MSE	MET	modified residue	UNP Q83K87
C	277	MSE	MET	modified residue	UNP Q83K87
C	300	MSE	MET	modified residue	UNP Q83K87
C	322	MSE	MET	modified residue	UNP Q83K87
D	-2	SER	-	cloning artifact	UNP Q83K87
D	-1	ASN	-	cloning artifact	UNP Q83K87
D	0	ALA	-	cloning artifact	UNP Q83K87
D	1	MSE	MET	modified residue	UNP Q83K87
D	58	MSE	MET	modified residue	UNP Q83K87
D	63	MSE	MET	modified residue	UNP Q83K87
D	69	MSE	MET	modified residue	UNP Q83K87
D	89	MSE	MET	modified residue	UNP Q83K87
D	123	MSE	MET	modified residue	UNP Q83K87
D	136	MSE	MET	modified residue	UNP Q83K87
D	161	MSE	MET	modified residue	UNP Q83K87
D	244	MSE	MET	modified residue	UNP Q83K87
D	277	MSE	MET	modified residue	UNP Q83K87
D	300	MSE	MET	modified residue	UNP Q83K87

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	322	MSE	MET	modified residue	UNP Q83K87
E	-2	SER	-	cloning artifact	UNP Q83K87
E	-1	ASN	-	cloning artifact	UNP Q83K87
E	0	ALA	-	cloning artifact	UNP Q83K87
E	1	MSE	MET	modified residue	UNP Q83K87
E	58	MSE	MET	modified residue	UNP Q83K87
E	63	MSE	MET	modified residue	UNP Q83K87
E	69	MSE	MET	modified residue	UNP Q83K87
E	89	MSE	MET	modified residue	UNP Q83K87
E	123	MSE	MET	modified residue	UNP Q83K87
E	136	MSE	MET	modified residue	UNP Q83K87
E	161	MSE	MET	modified residue	UNP Q83K87
E	244	MSE	MET	modified residue	UNP Q83K87
E	277	MSE	MET	modified residue	UNP Q83K87
E	300	MSE	MET	modified residue	UNP Q83K87
E	322	MSE	MET	modified residue	UNP Q83K87
F	-2	SER	-	cloning artifact	UNP Q83K87
F	-1	ASN	-	cloning artifact	UNP Q83K87
F	0	ALA	-	cloning artifact	UNP Q83K87
F	1	MSE	MET	modified residue	UNP Q83K87
F	58	MSE	MET	modified residue	UNP Q83K87
F	63	MSE	MET	modified residue	UNP Q83K87
F	69	MSE	MET	modified residue	UNP Q83K87
F	89	MSE	MET	modified residue	UNP Q83K87
F	123	MSE	MET	modified residue	UNP Q83K87
F	136	MSE	MET	modified residue	UNP Q83K87
F	161	MSE	MET	modified residue	UNP Q83K87
F	244	MSE	MET	modified residue	UNP Q83K87
F	277	MSE	MET	modified residue	UNP Q83K87
F	300	MSE	MET	modified residue	UNP Q83K87
F	322	MSE	MET	modified residue	UNP Q83K87

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total Zn 2 2	0	0
2	B	2	Total Zn 2 2	0	0
2	C	2	Total Zn 2 2	0	0
2	D	2	Total Zn 2 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	E	2	Total 2	Zn 2	0	0
2	F	2	Total 2	Zn 2	0	0

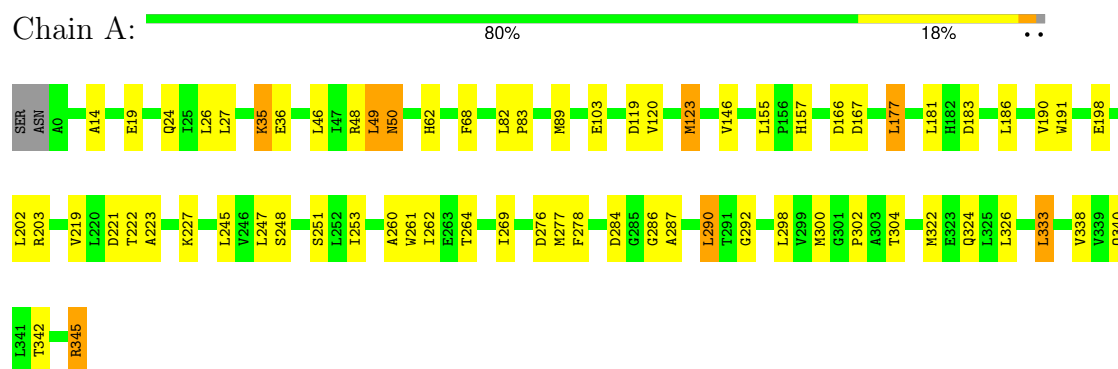
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	111	Total 111	O 111	0	0
3	B	114	Total 114	O 114	0	0
3	C	115	Total 115	O 115	0	0
3	D	120	Total 120	O 120	0	0
3	E	111	Total 111	O 111	0	0
3	F	106	Total 106	O 106	0	0

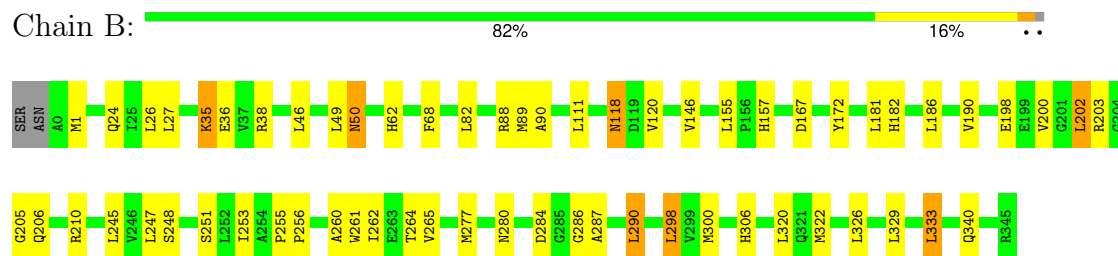
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

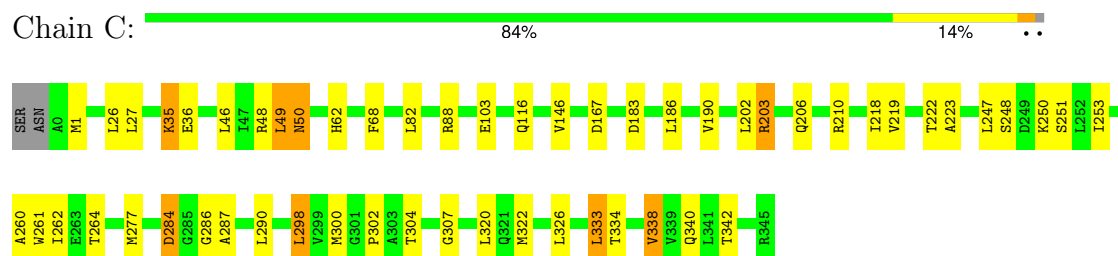
- Molecule 1: hypothetical protein SF2450



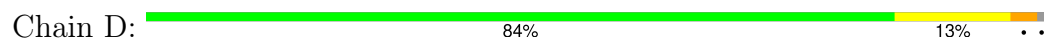
- Molecule 1: hypothetical protein SF2450



- Molecule 1: hypothetical protein SF2450



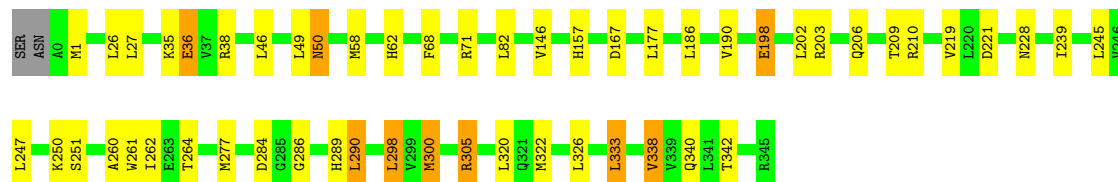
- Molecule 1: hypothetical protein SF2450





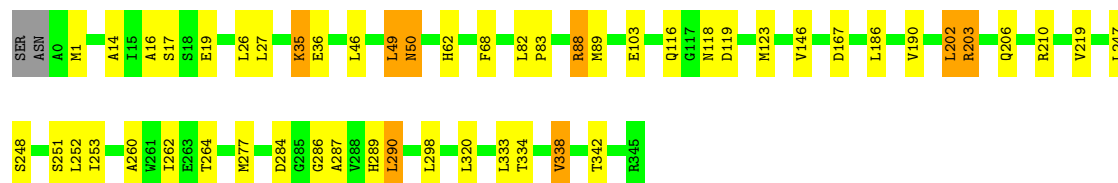
- Molecule 1: hypothetical protein SF2450

Chain E: 84% 13% ..



- Molecule 1: hypothetical protein SF2450

Chain F: 84% 13% ..



4 Data and refinement statistics

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, α , β , γ	161.83Å 161.83Å 228.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.15 50.00 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.00-2.15) 99.4 (50.00-2.15)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.95 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.190 , 0.229 0.200 , 0.235	Depositor DCC
R_{free} test set	7909 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å ²)	33.7	Xtriage
Anisotropy	0.035	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 19.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.467 for -1/2*h+1/2*k-1/2*l,1/2*h-1/2*k-1/2*l,-h-k 0.467 for -1/2*h+1/2*k+1/2*l,1/2*h-1/2*k+1/2*l,h+k 0.469 for -1/2*h-1/2*k+1/2*l,-1/2*h-1/2*k-1/2*l,h-k 0.469 for -1/2*h-1/2*k-1/2*l,-1/2*h-1/2*k+1/2*l,-h+k 0.477 for -h,k,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	16577	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section:
ZN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.59	0/2685	0.68	1/3619 (0.0%)
1	B	0.60	0/2685	0.67	0/3619
1	C	0.60	0/2685	0.67	0/3619
1	D	0.59	0/2685	0.67	1/3619 (0.0%)
1	E	0.58	0/2685	0.69	2/3619 (0.1%)
1	F	0.60	0/2685	0.67	0/3619
All	All	0.59	0/16110	0.67	4/21714 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	221	ASP	CB-CG-OD1	6.55	124.20	118.30
1	E	221	ASP	CB-CG-OD1	6.28	123.95	118.30
1	D	221	ASP	CB-CG-OD1	5.40	123.16	118.30
1	E	305	ARG	NE-CZ-NH1	5.17	122.88	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2648	0	2680	56	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2648	0	2680	42	0
1	C	2648	0	2680	43	0
1	D	2648	0	2680	41	0
1	E	2648	0	2680	40	0
1	F	2648	0	2680	39	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
2	E	2	0	0	0	0
2	F	2	0	0	0	0
3	A	111	0	0	3	0
3	B	114	0	0	2	0
3	C	115	0	0	2	0
3	D	120	0	0	3	0
3	E	111	0	0	4	0
3	F	106	0	0	3	0
All	All	16577	0	16080	245	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 8.

The worst 5 of 245 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:89:MSE:CE	1:A:120:VAL:H	1.52	1.22
1:E:206:GLN:HE21	1:E:210:ARG:HH12	1.08	1.02
1:B:206:GLN:HE21	1:B:210:ARG:HH12	1.04	0.96
1:A:222:THR:HG21	1:A:304:THR:HG23	1.48	0.96
1:D:1:MSE:HE3	1:D:320:LEU:CD2	1.97	0.93

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	348/348 (100%)	340 (98%)	8 (2%)	0	100	100
1	B	348/348 (100%)	342 (98%)	6 (2%)	0	100	100
1	C	348/348 (100%)	343 (99%)	5 (1%)	0	100	100
1	D	348/348 (100%)	342 (98%)	6 (2%)	0	100	100
1	E	348/348 (100%)	340 (98%)	8 (2%)	0	100	100
1	F	348/348 (100%)	342 (98%)	5 (1%)	1 (0%)	37	34
All	All	2088/2088 (100%)	2049 (98%)	38 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	203	ARG

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	286/272 (105%)	267 (93%)	19 (7%)	14	9
1	B	286/272 (105%)	267 (93%)	19 (7%)	14	9
1	C	286/272 (105%)	269 (94%)	17 (6%)	16	12
1	D	286/272 (105%)	269 (94%)	17 (6%)	16	12
1	E	286/272 (105%)	268 (94%)	18 (6%)	15	10
1	F	286/272 (105%)	272 (95%)	14 (5%)	21	18
All	All	1716/1632 (105%)	1612 (94%)	104 (6%)	16	12

5 of 104 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	35	LYS
1	D	338	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	202	LEU
1	D	49	LEU
1	D	247	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	206	GLN
1	F	50	ASN
1	F	182	HIS
1	C	50	ASN
1	B	206	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 12 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	334/348 (95%)	-0.98	0 100 100	18, 32, 54, 64	5 (1%)
1	B	334/348 (95%)	-0.97	0 100 100	19, 33, 54, 65	5 (1%)
1	C	334/348 (95%)	-0.95	0 100 100	18, 33, 54, 65	5 (1%)
1	D	334/348 (95%)	-0.96	0 100 100	17, 33, 53, 65	5 (1%)
1	E	334/348 (95%)	-0.95	0 100 100	18, 33, 54, 65	5 (1%)
1	F	334/348 (95%)	-0.99	0 100 100	18, 33, 54, 65	5 (1%)
All	All	2004/2088 (95%)	-0.96	0 100 100	17, 33, 54, 65	30 (1%)

There are no RSRZ outliers to report.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	ZN	D	1747	1/1	0.97	0.07	48,48,48,48	1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	ZN	C	1647	1/1	0.98	0.04	47,47,47,47	1
2	ZN	B	1547	1/1	0.99	0.03	49,49,49,49	1
2	ZN	C	1646	1/1	0.99	0.05	38,38,38,38	1
2	ZN	A	1447	1/1	0.99	0.03	47,47,47,47	1
2	ZN	D	1746	1/1	0.99	0.04	37,37,37,37	1
2	ZN	B	1546	1/1	0.99	0.08	36,36,36,36	1
2	ZN	E	1846	1/1	0.99	0.06	36,36,36,36	1
2	ZN	F	1946	1/1	0.99	0.04	36,36,36,36	1
2	ZN	E	1847	1/1	1.00	0.02	49,49,49,49	1
2	ZN	A	1446	1/1	1.00	0.05	37,37,37,37	1
2	ZN	F	1947	1/1	1.00	0.05	49,49,49,49	1

6.5 Other polymers [i](#)

There are no such residues in this entry.